

Maximum-likelihood astrometric geometry calibration of interferometric telescopes: application to the Very Small Array

Klaus Maisinger, M.P. Hobson, Richard D.E. Saunders and Keith J.B. Grainge

Astrophysics Group, Cavendish Laboratory, Madingley Road, Cambridge, CB3 0HE, UK

Accepted —. Received —; in original form 9 December 2002

ABSTRACT

Interferometers require accurate determination of the array configuration in order to produce reliable observations. A method is presented for finding the maximum-likelihood estimate of the telescope geometry, and of other instrumental parameters, astrometrically from the visibility timelines obtained from observations of celestial calibrator sources. The method copes systematically with complicated and unconventional antenna and array geometries, with electronic bandpasses that are different for each antenna radiometer, and with low signal-to-noise ratios for the calibrators. The technique automatically focusses on the geometry errors that are most significant for astronomical observation. We apply this method to observations made with the Very Small Array and constrain some 450 telescope parameters, such as the antenna positions, effective observing frequencies and correlator amplitudes and phase shifts; this requires only ~ 1 h of CPU time on a typical workstation.

Key words: instrumentation: interferometers – methods: observational – techniques: interferometric.

1 INTRODUCTION

Accurate knowledge of observing frequency, projected baselines and other geometric parameters is essential if interferometer data are to be turned into useful images. The accuracy required is typically a small fraction of the operating wavelength. For example, to keep phase errors below 10 degrees for an interferometer operating at 30 GHz, it is necessary to know the baseline length to an accuracy of better than 0.3 mm.

The standard way in which interferometer geometries are determined is along the following lines. First, direct measurements are made using, for example, tape measures, theodolites, and measurements of electrical lengths of electronic systems. Then, observations of celestial calibrators are carried out. For example, the observation by a particular pair of antennas of a bright calibrator of known RA and Dec gives the collimation phase of that pair (the difference in electrical length of the two arms of that interferometer). How this quantity varies, for example, with hour angle and for a series of calibrators at different declinations, is also investigated. These standard methods are described in e.g. Thompson, Moran & Swenson (1994), and in these ways a series of corrections to the directly measured geometry is produced, with the aim of ensuring that the physical baselines (in wavelengths) used in subsequent data analysis are correct.

This approach is of proven effectiveness for determining the geometries of interferometer arrays such as the Very Large Array (VLA). It is not adequate, however, for interferometers such as the Very Small Array (VSA, see e.g. Watson et al. (2002); Scott et al. (2002)), which is designed for imaging the CMB on scales $\gtrsim 10$

arcmin at 30 GHz, and thus has baselines of length $\lesssim 2$ m. At first sight this may seem surprising, as VSA baselines (in both metres and wavelengths) are so small. From the point of view of geometry determination from astrometry, however, the VSA is different from most other interferometer arrays in two respects.

- The collecting area of VSA antennas is very small and so there are few suitable celestial calibrators. Only the brightest calibrators are of use and signal-to-noise is a major issue.
- For most arrays like the VLA, the principal swivel axes of each antenna (e.g. altitude and azimuth, or polar and declination) intersect, at least nominally. If an axis is incorrectly aligned (e.g. the polar axis does not quite point to the Pole), then to first order this has no effect on the measured phase of a source, and its effect on source amplitude will manifest as a simple pointing correction which a standard 5-point observation (at expected source position, then offsets North, South, East then West) will provide. In the case of the VSA, the table elevation axis and the individual horn rotation axes (see Section 3) do not intersect by ~ 1 m. This design helps reduce important systematic errors and also reduces cost, but it makes astrometric geometry determination significantly harder. If an axis is misaligned, then in general a phase error is also introduced. The identification of such errors is exacerbated by the lack of signal-to-noise.

There are also further difficulties associated with determining the geometry of the VSA, which can be present in other interferometers.

- As a result of, for example, the compliance needed for fitting horn-reflectors to cryostats, the position of the electrical centre of

each horn will be slightly different, manifesting as a physical baseline correction and a pointing correction.

- A large observing bandwidth makes it hard to achieve a flat band pass, so the effective radio frequency becomes antenna dependent.
- Unconventional geometry makes it difficult to make a straightforward interpretation of a set of geometry calibration observations.

As a result of all these factors, phase (and amplitude) effects arise from complicated causes that are particularly hard to determine because of the lack of very high signal-to-noise calibrators. We have therefore implemented a maximum-likelihood method to determine the geometry, and other telescope parameters, astrometrically from the visibility timeline data obtained during a calibration observation of a sufficiently bright unresolved radio source. In the geometric model, we include as variables all the important dimensions and angles that require accurate determination, including the positions of the antennas, their effective observing frequencies and phase shifts on different baselines. In finding the best-fit geometry, the maximum-likelihood method correctly allows for the signal-to-noise ratio of the observations (on the assumption that the instrumental noise is Gaussian); the procedure also automatically guards against attempting to determine unimportant geometry errors that do not in fact much affect astronomical observation, since these similarly will not much affect calibrator observations.

In this paper, we focus on the method in its application to the VSA, although we point out the usefulness of the approach generally to radio and optical interferometry. In Section 2, we present the general maximum-likelihood method for the geometry calibration. The VSA is introduced in Section 3, where we develop a model of the telescope as a function of the parameters to be calibrated. Using this model, the likelihood method is applied to real VSA observations in Section 4. Finally, our conclusions are presented in Section 5.

2 ASTROMETRIC GEOMETRY CALIBRATION

Given a multi-dish interferometer consisting of N_a separate antennas, with antenna positions denoted by \mathbf{d}_i ($i = 1, \dots, N_a$) relative to some origin, one can measure the correlated signals of any two antennas along $N_b = \frac{1}{2}N_a(N_a - 1)$ independent baselines \mathbf{D}_j ($j = 1, \dots, N_b$), given in some ground-based reference frame specific to the telescope. The phase of the signal is modulated by the geometric path difference that the incident radiation experiences between the antennas. If the baseline is represented by a vector \mathbf{u} in a coordinate system (uv -coordinates) such that the component u of the vector $\mathbf{u} = (u, v, w)$ is measured in the east-west direction, v in the north-south direction and w in the direction of the antennas' pointing centre, the path difference is given by the projection of the direction cosines $\mathbf{x} = (x, y, z)$ of the source on the sky onto the baseline components, i.e. by $\Delta = \mathbf{u} \cdot \mathbf{x}$. Because the field of view is a section of a sphere, the third direction cosine is $z = \sqrt{1 - x^2 - y^2}$. Assuming a small field size or a sufficiently compact source, $z \approx 1$.

The response of an interferometer to an extended source can be obtained by integration over the solid angle of the source. Let us denote the sky brightness distribution of the source at the observing frequency ν by $I(x, y, \nu)$. One defines the complex (not fringe-rotated) visibility $V(u, v, w, \nu)$ as (Thompson 1994)

$$V(u, v, w, \nu) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(x, y, \nu) I(x, y, \nu) e^{2\pi i \frac{\nu}{c} (ux + vy + w)} dx dy, \quad (1)$$

where $A(x, y, \nu)$ is the directional antenna (power) response pattern

or *primary beam* (normalised to unity at its peak). To simplify the notation, we denote the real part $\Re[V(\mathbf{u})]$ of the visibility $V(\mathbf{u})$ by $\Re(\mathbf{u})$ and similarly $\Im(\mathbf{u})$ for the imaginary part.

For point sources at positions \mathbf{x}_k from the pointing centre, whose brightness distributions are given by $I(\mathbf{x}) = S_k \delta(\mathbf{x} - \mathbf{x}_k)$, the integral (1) for the complex visibilities $V(\mathbf{u}, \nu)$ reduces to

$$\begin{aligned} \Re(\mathbf{u}, \nu) &= \sum_k A(\mathbf{x}_k, \nu) S_k \cos\left(2\pi \frac{\nu}{c} \mathbf{u} \cdot \mathbf{x}_k\right), \\ \Im(\mathbf{u}, \nu) &= \sum_k A(\mathbf{x}_k, \nu) S_k \sin\left(2\pi \frac{\nu}{c} \mathbf{u} \cdot \mathbf{x}_k\right), \end{aligned} \quad (2)$$

where the sum is over all relevant sources in the field of view and \mathbf{x}_k and S_k denote, for the k th source, respectively the position from the pointing centre and the flux.

As the Earth rotates, the \mathbf{u} -vector corresponding to a baseline \mathbf{D}_j varies with the time t_i at which the visibility sample was measured. We will denote the uv -coordinates of the j th baseline at a time t_i (or alternatively at hour angle H_i) by $\mathbf{u}_{ij} \equiv \mathbf{u}(\mathbf{D}_j, t_i)$, and the corresponding visibility by V_{ij} . The transformation $\mathbf{R}(\delta, H_i, L)$ that relates a baseline vector \mathbf{D}_j in the telescope frame to the uv -coordinates is a simple rotation given by

$$\mathbf{u}_{ij} = \mathbf{R}(\delta, H_i, L) \mathbf{D}_j, \quad (3)$$

which will generally depend on the geographic latitude L of the telescope and the declination δ and hour angle H_i of the phase-tracking centre.

For a given a model of the telescope, the signal timelines for an observation of a point source can be predicted and compared with the observed data. If the source is sufficiently bright, one may then use a maximum-likelihood approach to determine the antenna geometry and a large number of other telescope parameters. For Gaussian instrumental noise the maximum-likelihood solution for the parameters can be found by minimising the standard χ^2 -misfit statistic between the predicted visibilities V^P and observed visibilities V^O , namely

$$\begin{aligned} \chi^2(V_{ij}^P, V_{ij}^O) &= \sum_{i,j}^{N_t, N_b} w_{ij} \left[(\Re^O(\mathbf{u}_{ij}) - \Re^P(\mathbf{u}_{ij}))^2 \right. \\ &\quad \left. + (\Im^O(\mathbf{u}_{ij}) - \Im^P(\mathbf{u}_{ij}))^2 \right]. \end{aligned} \quad (4)$$

In this expression, the sums run over all N_t time samples and N_b baselines. The weights on the observed visibilities V_{ij}^O are given by $w_{ij} = 1/\sigma_{ij}^2$, assuming that the noises are independent, which is true for the thermal Johnson noises from each antenna. For a given observation, χ^2 is a function of N_p parameters $\{p_l\}$ ($l = 1, \dots, N_p$), including, for example, the antenna positions on the table $\{\mathbf{d}_i\}$, the correlator amplitudes $\{C_j\}$, (constant) residual phase shifts $\{\phi_j\}$ and effective observing frequencies $\{\nu_i\}$. By minimising the χ^2 -function with respect to these parameters, it is possible to find the maximum-likelihood estimate of the telescope geometry.

Finding the global minimum of the χ^2 -function is numerically quite challenging. The optimisation takes place in an N_p -dimensional space and, as outlined for the VSA below, one typically has $N_p \approx 450$. Moreover, the $\chi^2 = \text{constant}$ hypersurfaces are rather uneven. In particular, there are many local minima, and a straightforward (e.g. gradient search) minimisation algorithm may never find the correct basin of attraction. The existence of local minima can be understood intuitively, since a minimum occurs whenever the fringe pattern is fitted correctly on only one of the baselines. Even though the false minimum can easily be recognised from the poor fit on at least some of the timelines, a straightforward

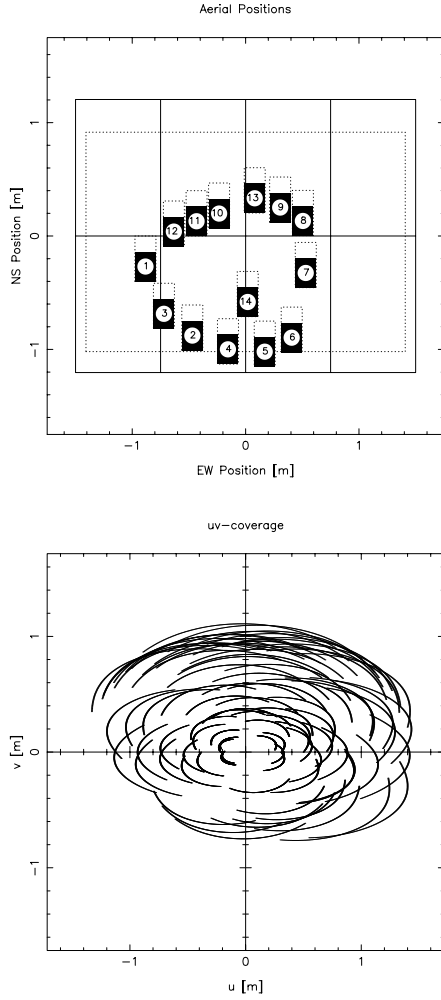


Figure 2. Top: the compact arrangement of the 14 VSA antennas on the table in September 2000. The rectangular area indicates the shape of the lower part of the VSA table, with the hinge on the lower end. The dashed lines above the horns show the extension of the antenna mount (Rusholme 2001). Bottom: the uv -coverage of the VSA using the antenna arrangement shown in (a). The observed field is at a declination of $\delta = 22^\circ$. The duration of the observation is 3 hours either side of transit.

j th baseline. The path delays are introduced at the different IF frequencies $\nu_{\text{IF},j,1}^{(r)}$ and $\nu_{\text{IF},j,2}^{(r)}$. Note that this notation with $2 \times N_b$ different symbols obscures the fact that there are just $2 \times N_a$ different path delays and IF frequencies, as the path delays are inserted per antenna and not per baseline.

Gathering together the effects described above, the predicted visibilities for the j th baseline for observations of a set of point sources at positions \mathbf{x}_k with fluxes S_k are thus given by

$$\begin{aligned} \Re^P(\mathbf{u}_j) &= C_j^{(r)} \sum_k A \left(\mathbf{x}_k, \sqrt{\nu_{j,1}^{(r)} \nu_{j,2}^{(r)}} \right) S_k \\ &\times \cos \left(2\pi \left[\frac{\sqrt{\nu_{j,1}^{(r)} \nu_{j,2}^{(r)}}}{c} \mathbf{u}_j \cdot \mathbf{x}_k \right. \right. \\ &\quad \left. \left. - \left\{ \frac{\nu_{\text{IF},j,1}^{(r)}}{c} \rho_{j,1}^{(r)} - \frac{\nu_{\text{IF},j,2}^{(r)}}{c} \rho_{j,2}^{(r)} \right\} + \phi_j^{(r)} \right] \right), \end{aligned}$$

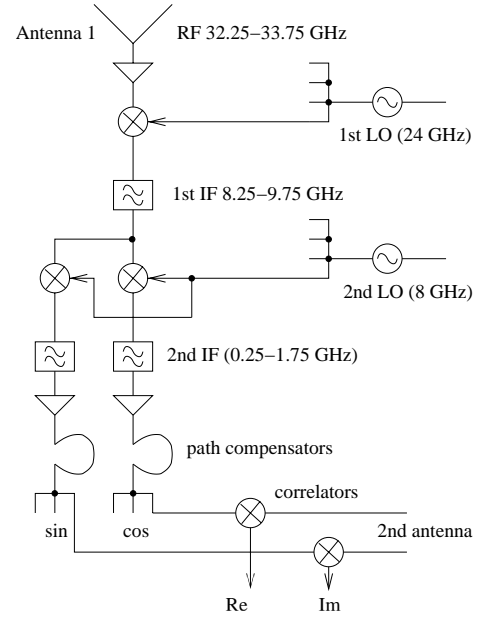


Figure 3. Simplified schematic of the electronics of the receiving system of the VSA for compact array observations at a centre frequency of 33 GHz. A signal coming from an antenna at the radio frequency ν_{RF} is reduced to the (2nd) IF frequency ν_{IF} after mixing with the signals coming from the two local oscillators. It is split into sin- and cos-channels, and a geometric path delay is inserted into each channel separately. The correlated signals from pairs of antennas are integrated to produce the real and imaginary parts of the visibility.

$$\begin{aligned} \Im^P(\mathbf{u}_j) &= C_j^{(i)} \sum_k A \left(\mathbf{x}_k, \sqrt{\nu_{j,1}^{(i)} \nu_{j,2}^{(i)}} \right) S_k \\ &\times \sin \left(2\pi \left[\frac{\sqrt{\nu_{j,1}^{(i)} \nu_{j,2}^{(i)}}}{c} \mathbf{u}_j \cdot \mathbf{x}_k \right. \right. \\ &\quad \left. \left. - \left\{ \frac{\nu_{\text{IF},j,1}^{(i)}}{c} \rho_{j,1}^{(i)} - \frac{\nu_{\text{IF},j,2}^{(i)}}{c} \rho_{j,2}^{(i)} \right\} + \phi_j^{(i)} \right] \right), \quad (6) \end{aligned}$$

where $\nu = \nu_{\text{IF}} + \nu_{\text{LO},1} + \nu_{\text{LO},2}$ is the radio frequency corresponding to a given IF frequency ν_{IF} . The superscripts P indicate that the visibilities have been theoretically predicted. Given the calculable dependence of \mathbf{u}_j for each baseline on the time t , one can thus calculate the quantities V_{ij}^P appearing in the expression for χ^2 given in (4).

4 ANALYSIS OF VSA CALIBRATION OBSERVATIONS

In order to constrain the geometry, an observation with a good signal-to-noise ratio at a sufficiently high elevation is required. There are three possible calibration sources for the VSA compact array that fulfil these requirements: Cas-A, Cyg-A and Tau-A. These are essentially unresolved point sources for the VSA compact array. In the plots presented below, we use a 4-h calibration observation of Cas-A as an illustrative example. The optimisation with respect to antenna positions, correlator amplitudes and phases and the IF frequencies was carried out using Powell's algorithm down to a relative convergence tolerance of 10^{-3} in the value of χ^2 . The entire calculation required around 1 h of CPU time on a Sparc Ultra workstation.

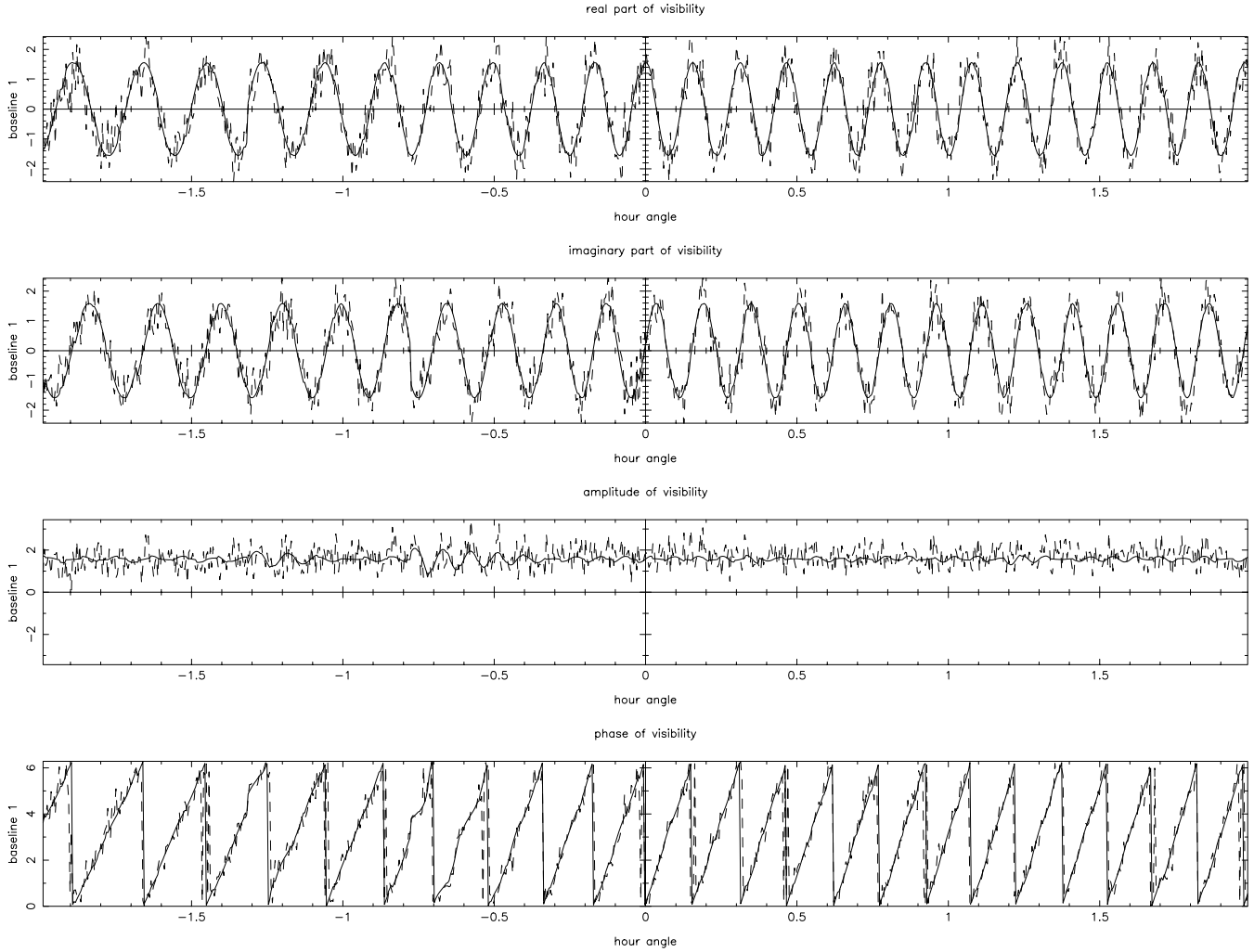


Figure 4. Comparison of observed and predicted fringes on VSA baseline antenna 1/antenna 2, using an observation of Cas-A. The panels show the real and imaginary parts, the amplitude and the phase of the visibility respectively, as a function of hour angle (in hours). The theoretical fringe predicted from the optimised telescope parameters is plotted with a solid line and the observed fringe with a dashed line. The units on the y-axis are arbitrary.

4.1 Timelines

As an illustration of the fit obtained, in Fig. 4 we plot the observed fringes for VSA baseline antenna 1/antenna 2. The four panels show respectively the real and imaginary parts of the visibility, the amplitude and the phase, each as a function of hour angle. Each panel shows the timelines for the observed fringes (dashed) and the fringes predicted assuming the best-fit telescope geometry (solid), obtained from the χ^2 -minimisation. The scatter of the observed real and imaginary parts around the predicted lines indicates the noise level of the observation.

The quasi-sinusoidal fringe pattern shows an increasing fringe frequency for the 1/2 baseline. Since the data have not been fringe-rotated, the phase is going rapidly through cycles of 2π . On this baseline, small quadrature errors (see Watson et al. (2002)) and differences in the amplitudes on the two channels are seen as a slight variation of the visibility amplitude and as structure in the visibility phase. We note that the best fit is in excellent agreement with the observations and matches the observed amplitudes and phases very closely.

4.2 The χ^2 -surface

Confirmation that the optimisation procedure has indeed fully converged can be seen in Fig. 5, 6 and 7. These plots show one-dimensional cuts through the χ^2 -surface in parameter space along some of the parameter coordinate axes. In each case, the y-axis indicates the value of the χ^2 -function, with an arbitrary normalisation, as a function of the deviation of the parameter value from the best-fit value. All plots are centred around a local minimum in the χ^2 -surface, indicating that the optimisation procedure has fully converged, and that the obtained reconstruction is indeed the maximum-likelihood estimate. There is, however, a strong caveat in interpreting these one-dimensional χ^2 -contours: they are by no means representative of the accuracy with which the parameters can be determined, which would require a marginalisation over all other parameters rather than a simple cut. In fact, the position and shape of the one-dimensional contour can be changed significantly by simply varying one or a few of the other parameters. To ensure that a minimum has been found, all parameters have to converge simultaneously.

Fig. 5 shows a cut along the x -position of VSA antenna 1.

Arg. 4: x pos ant 1

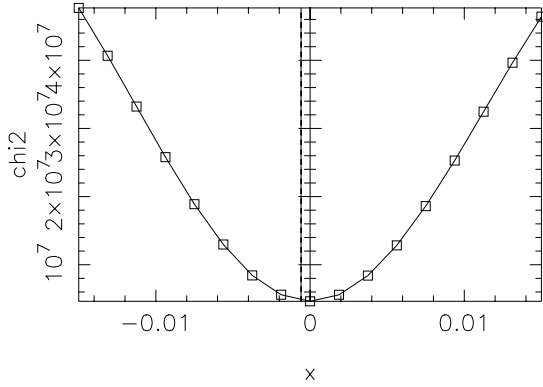


Figure 5. The χ^2 -surface as a function of the x -position of VSA antenna 1. The units on the x -axis are metres from the optimised position, the y -units are arbitrary. Low values of χ^2 correspond to the most likely solution. The minimisation has converged to the local minimum in the centre. A vertical line indicates the position of the manually measured value of the position.

Arg. 18: y pos ant 1

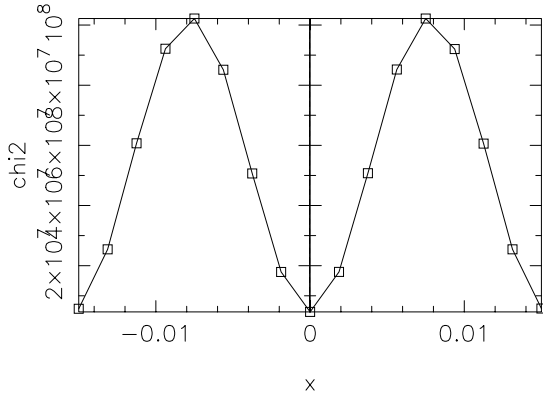


Figure 6. The χ^2 -surface as a function of the y -position of antenna 1. Compare Fig. 5. It is interesting to note that adjacent minima are nearly as pronounced as the central one.

Arg. 410: cos frequency ant 1

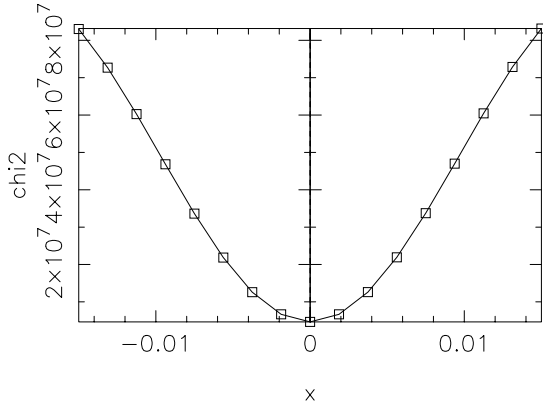


Figure 7. The χ^2 -surface as a function of the IF frequency on the cos-channel of antenna 1. The x -axis is in units of 12 GHz from the optimal frequency, i.e. 0.01 corresponds to 120 MHz.

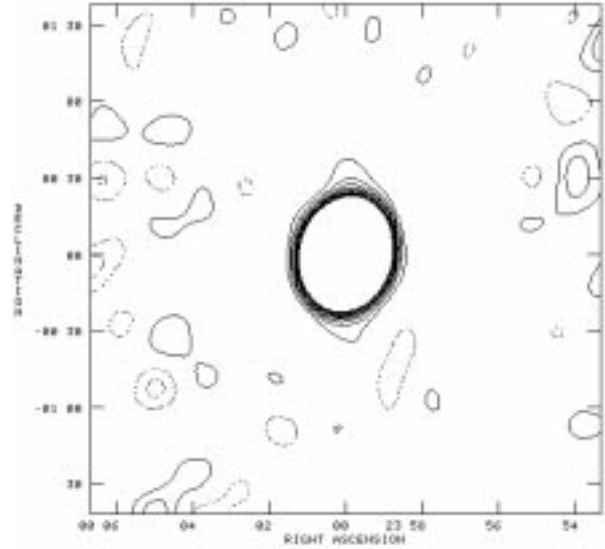


Figure 8. An image of Jupiter observed by the VSA. The geometry and other telescope parameters have been simultaneously calibrated by observations of Tau-A and Cas-A. The dynamical range of the image is 500:1, indicating that the phase errors due to the calibration are indeed negligible.

Minima in χ^2 correspond to the most likely solution. As for all other plots, the minimum has been found. A vertical line indicates the original input value of the position, as was manually measured on the table. Fig. 6 shows a cut along the y -position for the same antenna. It is obvious from the figure that, in this case, the calibration observation used is only just sufficient to distinguish between the true minimum and the adjacent minima either side, which are nearly as deep. As mentioned previously, to avoid being trapped in a false local minimum, one requires a sufficiently accurate initial estimate. Nevertheless, it is still possible to detect false minima, since the resulting positioning mistake of the order of 1 cm induces significant phase errors on at least one of the baselines of which the corresponding antenna is a part.

As a further illustration, Fig. 7 shows a cut along the IF frequency of the cos-channel of antenna 1. Once more, we see that the position obtained by the Powell algorithm lies indeed at the minimum for this parameter.

4.3 Imaging with the VSA

The parameters found from the calibration observations can be applied to the data reduction of subsequent observations. Fig. 8 shows a VSA map of Jupiter, which has been produced using a calibration that combined observations of both Cas-A and Tau-A. This highlights that the geometry fitting procedure can easily exploit several data sets simultaneously by simply adding up the χ^2 -functionals for all observations. The dynamical range of the image is an excellent 500:1, showing that the geometry calibration on other sources was sufficiently successful to render phase errors negligible. Other VSA maps are shown in Watson et al. (2002), and the primordial CMB maps obtained from VSA observations in the compact configuration are presented by Taylor et al. (2002).

5 DISCUSSION AND CONCLUSIONS

In this paper, we have presented a method for astrometric geometry calibration of interferometers by performing a likelihood analysis of visibility timelines produced during an observation of a reasonably bright radio point source. The method has been demonstrated for the specific case of the Very Small Array, for which one must simultaneously constrain some 450 telescope parameters, such as the antenna positions, the effective observing frequency of each antenna which can vary due to different bandpass characteristics, and phase shifts on different baselines introduced in the electronic system or by different cable lengths. In particular, analysis of VSA observations of Tau-A and Cas-A sources show that the relative positions of the antennas on the telescope table can be determined to a small fraction of a wavelength. The quality of the fit is sufficiently good to enable high dynamic-range mapping with the VSA. This astrometric calibration is very straightforward and cheap, as it requires no equipment other than the telescope itself. Multiple calibration observations can easily be combined to produce a better parameter estimate by summing the χ^2 -functionals for all available data sets.

The calibration method is currently in active use by the VSA project and has successfully determined the geometry for several different antenna configuration since the start of its observing programme. The calibration of the VSA is described in more detail in Watson et al. (2002), and power spectra derived from the first year of observations in the compact configuration are presented in Scott et al. (2002).

Although the astrometric calibration method has been applied specifically to the VSA, it can easily be adapted to other telescopes. It is worth noting, however, that there is one feature of the VSA which does not necessarily generalise to other CMB telescopes, even though it is common to most ground-based arrays. Each VSA antenna tracks the source separately in ‘azimuth’ (*semi-comounted* design). This has the advantage of providing a valuable check on systematic contaminations, because the path delays on a particular baseline change during the observation and the sky signal can be recognised by its characteristic fringe frequency. On the Cosmic Background Imager (Padin et al. 2002) and the Degree Angular Scale Interferometer (Leitch et al. 2002), the antennas are mounted on a rigid tracking platform supported by an alt-azimuth mount. This design keeps the baseline projections with respect to the source constant and reduces the fringe rate on all baselines to zero, making the described geometry fitting impossible. A similar method can still be applied, however, if the fringe rate is artificially changed. For instance, the telescope could be kept at a fixed pointing position while the source is within the antenna beam, thus providing a non-zero fringe rate.

ACKNOWLEDGMENTS

We thank Mike Jones, Guy Pooley, Anze Slosar and many other members of the VSA team for their helpful discussions. Klaus Maisinger acknowledges support from an EU Marie Curie Fellowship.

REFERENCES

- Kirkpatrick S., Gelatt Jr. C.D., Vecchi M.P., 1983, *Science*, 220, 671
- Leitch E.M., 2002, *ApJ*, 568, 28

- Padin S. et al., 2002, *PASP*, 114, 83
- Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P., 1992, *Numerical Recipies in Fortran*. CUP, Cambridge
- Rusholme B., 2001, *The Very Small Array*. PhD thesis, University of Cambridge
- Scott P.F. et al., 2002, submitted to *MNRAS*, astro-ph/0205380
- Taylor A.C. et al., 2002, submitted to *MNRAS*, astro-ph/0205381
- Thompson A.R., Moran J.M., Swenson G.W., 1994, *Interferometry and Synthesis in Radio Astronomy*. Krieger Publishing Company, New York
- Watson R.A. et al., 2002, submitted to *MNRAS*, astro-ph/0205378

APPENDIX A: DERIVATIVES USED IN THE VSA GEOMETRY CALIBRATION

For reference purposes, in this Appendix we quote the expressions for the derivatives of the χ^2 -functional (4) combined with the visibilities (6) used in the maximum-likelihood algorithm for the the VSA geometry calibration.

The gradient is required by some numerical minimisation algorithms that use gradient information to accelerate the computation. Although not required by the minimisation algorithm, the curvature can be useful for some iterative minimisation algorithms and for quantifying errors or uncertainties on the reconstruction. Assuming that the maximised estimate of the parameters is indeed in the correct basin of attraction on the likelihood surface and that this surface can be reasonably approximated by a Gaussian around the maximum, the uncertainties in the reconstruction can be obtained from the inverse of the Hessian matrix of χ^2 .

A1 First derivatives of χ^2

From (4), the χ^2 -gradient with respect to the telescope and correlator parameters p_l is

$$\frac{\partial \chi^2}{\partial p_l} = 2 \sum_{i,j}^{N_t, N_b} w_{ij} \left[\left(\Re^P(u_{ij}) - \Re^O(u_{ij}) \right) \frac{d\Re^P(u_{ij})}{dp_l} + \left(\Im^P(u_{ij}) - \Im^O(u_{ij}) \right) \frac{d\Im^P(u_{ij})}{dp_l} \right]$$

where again the symbol $\Re^P(u_{ij})$ denotes the real part of the visibility predicted for the i th time sample on the j th baseline. The dependence of the visibilities on the parameters can be explicit, as in the case of the correlator amplitudes, or implicit in the position of the samples u_{ij} , as for the antenna positions:

$$\frac{d\Re^P(u_{ij})}{dp_l} = \frac{\partial \Re^P(u_{ij})}{\partial p_l} + \frac{\partial \Re^P(u_{ij})}{\partial u_{ij}} \cdot \frac{\partial u_{ij}}{\partial p_l}. \quad (\text{A1})$$

We now substitute the expressions (6) appropriate for the VSA. For the sake of notational brevity, we define the arguments

$$\psi^{(r)} = \frac{\sqrt{v_1^{(r)} v_2^{(r)}}}{c} \mathbf{u}_j \cdot \mathbf{x}_k - \left\{ \frac{v_{\text{IF},1}^{(r)}}{c} \rho_1^{(r)} - \frac{v_{\text{IF},2}^{(r)}}{c} \rho_2^{(r)} \right\} + \phi^{(r)},$$

$$v^{(r)} = \sqrt{v_1^{(r)} v_2^{(r)}}, \quad (\text{A2})$$

and similarly for $\psi^{(i)}$ and $v^{(i)}$. Then the derivatives of the predicted visibilities with respect to the uv -loci,

$$\begin{aligned} \frac{\partial \Re^P(\mathbf{u})}{\partial \mathbf{u}} &= -2\pi C^{(r)} \sum_k A(\mathbf{x}_k, v^{(r)}) S_k \sin(2\pi \psi^{(r)}) \frac{v^{(r)}}{c} \mathbf{x}_k, \\ \frac{\partial \Im^P(\mathbf{u})}{\partial \mathbf{u}} &= 2\pi C^{(i)} \sum_k A(\mathbf{x}_k, v^{(i)}) S_k \cos(2\pi \psi^{(i)}) \frac{v^{(i)}}{c} \mathbf{x}_k, \end{aligned} \quad (\text{A3})$$

only depend on table parameters, whereas

$$\begin{aligned} \frac{\partial \Re^P(\mathbf{u})}{\partial C^{(r)}} &= \sum_k A(\mathbf{x}_k, v^{(r)}) S_k \cos(2\pi \psi^{(r)}), \\ \frac{\partial \Im^P(\mathbf{u})}{\partial C^{(i)}} &= \sum_k A(\mathbf{x}_k, v^{(i)}) S_k \sin(2\pi \psi^{(i)}), \end{aligned} \quad (\text{A4})$$

$$\begin{aligned} \frac{\partial \Re^P(\mathbf{u})}{\partial \phi^{(r)}} &= -2\pi C^{(r)} \sum_k A(\mathbf{x}_k, v^{(r)}) S_k \sin(2\pi \psi^{(r)}), \\ \frac{\partial \Im^P(\mathbf{u})}{\partial \phi^{(i)}} &= 2\pi C^{(i)} \sum_k A(\mathbf{x}_k, v^{(i)}) S_k \cos(2\pi \psi^{(i)}), \end{aligned} \quad (\text{A5})$$

only depend on parameters of the electronic system.

The derivatives with respect to the frequencies are

$$\begin{aligned} \frac{\partial \Re^P(\mathbf{u})}{\partial v_{\text{IF},1}^{(r)}} &= -2\pi C^{(r)} \sum_k A(\mathbf{x}_k, v^{(r)}) S_k \sin(2\pi \psi^{(r)}) \left(\frac{1}{2c} \sqrt{\frac{v_2^{(r)}}{v_1^{(r)}}} - \frac{\rho_1^{(r)}}{c} \right) \mathbf{u}_j \cdot \mathbf{x}_k, \\ \frac{\partial \Re^P(\mathbf{u})}{\partial v_{\text{IF},2}^{(r)}} &= -2\pi C^{(r)} \sum_k A(\mathbf{x}_k, v^{(r)}) S_k \sin(2\pi \psi^{(r)}) \left(\frac{1}{2c} \sqrt{\frac{v_1^{(r)}}{v_2^{(r)}}} + \frac{\rho_2^{(r)}}{c} \right) \mathbf{u}_j \cdot \mathbf{x}_k, \end{aligned} \quad (\text{A6})$$

and analogously for the imaginary parts. These derivatives apply to a single baseline only. In practice, a given frequency channel is used in several baselines. In order to compute the χ^2 -derivatives one thus has to sum over all baselines containing the respective frequency channel.

The derivatives of the uv -positions with respect to the geometric parameters are given by

$$\frac{\partial(u_{ij})_k}{\partial p_l} = \begin{cases} \sum_m \frac{\partial R_{km}^i}{\partial p_l} D_m^j & \text{if } p_l \in \{\Delta\epsilon, \theta, \phi\} \\ \sum_m R_{km}^i \frac{\partial D_m^j}{\partial p_l} & \text{if } p_l \in \{\mathbf{d}_i\} \\ 0 & \text{otherwise} \end{cases}, \quad (\text{A7})$$

where subscripts denote the vector components, \mathbf{D} is a baseline vector and $\mathbf{R}(\delta, H_i, L, \Delta\epsilon, \theta, \chi)$ is the transformation matrix from table to uv -coordinates introduced in (5):

$$\mathbf{u}_{ij} = \mathbf{R}(\delta, H, L, \Delta\epsilon, \theta, \chi) \mathbf{D}_j(\{\mathbf{d}_i\}).$$

Derivatives with respect to parameters other than the table parameters will vanish since they do not affect the geometry. Furthermore, a given baseline depends only on the positions of those two horns of which it is composed, and the resulting matrix is considerably sparse.

A2 Curvature of χ^2

Denoting the real part of the visibility predicted at the i th time sample on baseline j by \Re_{ij}^P , the χ^2 -curvature is given by

$$\begin{aligned} \frac{\partial^2 \chi^2}{\partial p_l \partial p_n} = & 2 \sum_{i,j} w_{ij} \left\{ \left(\frac{\partial \Re_{ij}^P}{\partial \mathbf{u}} \cdot \frac{\partial \mathbf{u}_{ij}}{\partial p_n} \right) \left(\frac{\partial \Re_{ij}^P}{\partial p_l} + \frac{\partial \Re_{ij}^P}{\partial \mathbf{u}} \cdot \frac{\partial \mathbf{u}_{ij}}{\partial p_l} \right) \right. \\ & + \left(\Re_{ij}^P - \Re_{ij}^O \right) \left[\frac{\partial^2 \Re_{ij}^P}{\partial p_l \partial p_n} + \left(\frac{\partial \mathbf{u}_{ij}}{\partial p_n} \right)^t \left[\nabla_{\mathbf{u}} \nabla_{\mathbf{u}} \Re_{ij}^P \right] \left(\frac{\partial \mathbf{u}_{ij}}{\partial p_l} \right) + \frac{\partial \Re_{ij}^P}{\partial \mathbf{u}} \cdot \frac{\partial^2 \mathbf{u}_{ij}}{\partial p_l \partial p_n} \right] \\ & + \left(\frac{\partial \Im_{ij}^P}{\partial \mathbf{u}} \cdot \frac{\partial \mathbf{u}_{ij}}{\partial p_n} \right) \left(\frac{\partial \Im_{ij}^P}{\partial p_l} + \frac{\partial \Im_{ij}^P}{\partial \mathbf{u}} \cdot \frac{\partial \mathbf{u}_{ij}}{\partial p_l} \right) \\ & \left. + \left(\Im_{ij}^P - \Im_{ij}^O \right) \left[\frac{\partial^2 \Im_{ij}^P}{\partial p_l \partial p_n} + \left(\frac{\partial \mathbf{u}_{ij}}{\partial p_n} \right)^t \left[\nabla_{\mathbf{u}} \nabla_{\mathbf{u}} \Im_{ij}^P \right] \left(\frac{\partial \mathbf{u}_{ij}}{\partial p_l} \right) + \frac{\partial \Im_{ij}^P}{\partial \mathbf{u}} \cdot \frac{\partial^2 \mathbf{u}_{ij}}{\partial p_l \partial p_n} \right] \right\}. \end{aligned}$$

Again the expressions (A3)–(A6) can be used to substitute for the partial derivatives.